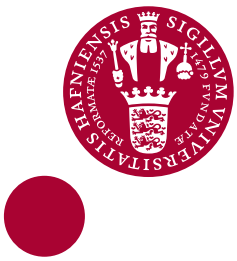

Evolutionary history of stomach bot flies in the light of mitogenomics

Yan, Liping; Pape, Thomas; Elgar, Mark A.; Gao, Yunyun; Zhang, Dong

Published in:
Systematic Entomology

DOI:
10.1111/syen.12356

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Systematic Entomology

Systematic Entomology (2019), **44**, 797–809 DOI: 10.1111/syen.12356

Evolutionary history of stomach bot flies in the light of mitogenomics

LIPING YAN[1], THOMAS PAPE[2], MARK A. ELGAR[3], YUNYUN GAO[1] and DONG ZHANG[1]

[1]School of Nature Conservation, Beijing Forestry University, Beijing, China, [2]Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark and [3]School of BioSciences, University of Melbourne, Melbourne, Australia

Abstract. Stomach bot flies (Calyptratae: Oestridae, Gasterophilinae) are obligate endoparasitoids of Proboscidea (i.e. elephants), Rhinocerotidae (i.e. rhinos) and Equidae (i.e. horses and zebras, etc.), with their larvae developing in the digestive tract of hosts with very strong host specificity. They represent an extremely unusual diversity among dipteran, or even insect parasites in general, and therefore provide significant insights into the evolution of parasitism. The phylogeny of stomach bot flies was reconstructed based on extensive mitochondrial genomic data for *Cobboldia*, *Gyrostigma* and six of the eight known species of *Gasterophilus*. The phylogenetic tree, i.e. {*Cobboldia*, [*Gyrostigma*, (*Gasterophilus pecorum*, (*Gasterophilus intestinalis*, (*Gasterophilus haemorrhoidalis*, *Gasterophilus inermis*)), (*Gasterophilus nasalis*, *Gasterophilus nigricornis*))]}, provides a strong evolutionary reference to infer several biological patterns for the first time for this group: (i) host shifts of stomach bot flies from elephants to rhinoceroses and then from rhinoceroses to equids; (ii) dispersal with their hosts from the Afrotropical region into the Palaearctic and Oriental regions; (iii) oviposition site, originally on the host head, and egg production positively correlated with distance from the mouth; (iv) attachment of third-instar larva originally in the stomach, with duodenal and large intestinal positions secondarily derived; and (v) guanine and cytosine enrichment of the mitogenome as an adaptation to larval life in the warm environment of the host digestive tract, combined with the need for a high evolutionary rate to cope with the fast evolution of their mammalian hosts.

Introduction

Parasites sensu lato comprise nearly half of animalian diversity (De Meeûs & Renaud, 2002), with parasitism estimated to have evolved independently at least 223 times (Weinstein & Kuris, 2016). Parasite life cycles typically comprise three stages: transmission, infection and establishment (Read, 1972; Zelmer, 1998), and each may require specific adaptations, depending upon the degree of host specificity of the parasite. Transmission to hosts may depend on host-specific oviposition strategies (e.g. Stireman *et al.*, 2006); infection may require mechanisms that ensure the parasite progeny to enter, attach to or associate with the host (e.g. Nufio & Papaj, 2004); and establishment requires the parasite progeny to locate and remain in particular sites within, on or associated with the host. These three properties of parasitism, which may be highly host-specific, make host shifts remarkable events of considerable evolutionary interest (De Fine Licht, 2018), especially when they result in speciation [e.g. nudibranchs (Faucci *et al.*, 2007), and parasitic flatworms (Zietara & Lumme, 2002)].

Parasitism has evolved numerous times within Diptera (Grimaldi & Engel, 2005; Courtney *et al.*, 2017), and more frequently than within any other group of insects (Feener Jr & Brown, 1997; Wiegmann *et al.*, 2011), and even of animals (Weinstein & Kuris, 2016). Bot flies (Oestridae) are obligate larval endoparasites of mammals (Colwell, 2006; Guimarães & Papavero, 1999; Zumpt, 1965), which is an unusual lifestyle among Diptera, with a possible early or mid-Eocene origin (Pape, 2006; Cerretti *et al.*, 2017; Stireman *et al.*, 2019). Bot flies may be the first dipteran group to have evolved mammal myiasis and, with approximately 170 extant species (Pape *et al.*,

Correspondence: Dong Zhang, School of Nature Conservation, Beijing Forestry University, Beijing, China.
E-mail: zhangdong_bjfu@bjfu.edu.cn; ernest8445@163.com

Table 1. Life-history parameters in stomach bot flies (see File S4 for literature sources).

Species	Host	Oviposition site	Fecundity	Attachment organ (AO) type	Infecting strategy of first-instar larvae	Third-instar attachment location
Cobboldia loxodontis	Elephantidae	Base of the tusks	Unknown	Unknown	Unknown	Stomach
Gyrostigma rhinocerontis	Rhinocerotidae	Mainly on the head, neck and shoulders	~750	Type II AO	Unknown	Stomach
Gasterophilus haemorrhoidalis	Equidae	Lips, mainly the upper lips	50–200	Type I AO	Hatching with stimulation of moisture from host, and spontaneously migrating into the host's mouth	Large intestine
Gasterophilus inermis	Equidae	Cheeks	320–360	Type I AO	Spontaneously hatching and migrating into the host's mouth	Large intestine
Gasterophilus intestinalis	Equidae	Mainly lower forelegs, also on the back and flanks	400–1000	Type I AO	Hatching and entering the host's mouth when licked by host	Stomach
Gasterophilus nasalis	Equidae	Under the chin in the groove between the halves of the lower jaw	300–500	Type I AO	Spontaneously hatching and migrating into the host's mouth	Pylorus and duodenum
Gasterophilus nigricornis	Equidae	On the cheeks, rarely on the nasal region	330–350	Type I AO	Spontaneously hatching and migrating into the host's mouth	Pylorus and duodenum
Gasterophilus meridionalis	Equidae	Unknown	Unknown	Type I AO	Unknown	Pylorus and duodenum
Gasterophilus pecorum	Equidae	Leaves and stems of plants, mainly grasses	1300–2500	Type II AO	Hatching and entering the host's mouth when eaten by host	Pharynx and stomach
Gasterophilus ternicinctus	Equidae	Unknown	Unknown	Type I AO	Unknown	Stomach

2011, 2017), are by far the largest radiation of dipteran mammal endoparasites. Although conventionally defined as endoparasites, the transmission biology of bot flies, with free-living adults, more closely resembles that of insect parasitoids, and so they may be more usefully referred to as mammal endoparasitoids (Zelmer, 1998) even if they do not kill their host. The larvae of most bot flies are subcutaneous parasitoids, but stomach bot flies develop in the digestive tract of their hosts with very strong host specificity (Colwell, 2006; Zumpt, 1965). The biology of this small group of bot flies is extremely unusual not only within Diptera, but also across the entire Insecta (Balashov, 2006).

The stomach bot flies contain three genera, *Cobboldia* Brauer, *Gyrostigma* Brauer and *Gasterophilus* Leach, which are parasitoids of Proboscidea (i.e. elephants), Rhinocerotidae (i.e. rhinos) and Equidae (i.e. horses, zebras, etc.) (Zumpt, 1965), respectively. Females oviposit on different areas of the host body (with one exception where eggs are laid on the food plants), and the first-instar larvae (LI) may hatch spontaneously or following host stimulation, subsequently entering the digestive tract through the mouth. Larval development lasts up to 10 months, with the mature third-instar larvae (LIII) leaving the host with the faeces via the anus. After leaving the host, the larvae pupate in the ground, eventually eclosing as adults to mate and oviposit. All stomach bot fly larvae attach to the wall of the digestive tract of their hosts, but the specific location varies among species. For example, there is evidence that the larvae of *Cobboldia* attach to the elephant stomach wall (Gowda *et al.*, 2017); the second-instar larvae (LII) and LIII of *Gyrostigma* attach to the stomach wall of their rhinoceros host (Zumpt, 1965); while the LIII of all *Gasterophilus* species attach to different locations along the digestive tract of their equine host (Horak *et al.*, 1984; Zumpt, 1965). Detailed life-history information for each stomach bot fly species is summarized in Table 1. With 14 known extant species (three species of *Cobboldia*, three of *Gyrostigma*, and eight of *Gasterophilus*), the stomach bot flies have a modest species richness (Pape, 2006), but this should be considered in the context of the constraints of their specialized niche, and many species are likely to have become extinct with their host, such as *Cobboldia russanovi* Grunin found from the woolly mammoth (Grunin, 1973).

We used data from the mitochondrial genome, which plays an important role in systematic research (e.g. Barker, 2014; Cameron, 2014; Timmermans *et al.*, 2014; Zhang *et al.*, 2016a; Horreo, 2017), in order to reconstruct the phylogeny of the stomach bot flies. We used this phylogeny as an evolutionary

framework to explain the life-history strategies as a result of diversification through shifts in hosts, oviposition sites, female fecundity and larval attachment sites, within this highly specialised group of remarkable parasitoids.

Materials and methods

DNA extraction, sequencing and annotation

DNA was extracted from LIII or adults preserved in 99.5% ethanol (File S1) following the protocol described in Zhang *et al.* (2016a) and stored at −20°C dissolved in Tris-EDTA buffer until use. Mitogenomic data for *Cobboldia* and *Gasterophilus* were amplified using primer pairs following PCR protocols in File S2. The PCR reaction, amplicon sequencing and fragment assembly were performed as described in Zhang *et al.* (2016a). Only the gene for cytochrome *c* oxidase subunit 2 (COII) and a part of the gene for subunit 1 (COI) were successfully amplified and sequenced for *Cobboldia loxodontis* (Brauer) because of limited amounts of DNA extracted from a single leg of the most recently collected, available specimen (File S1). The mitogenome of *Gyrostigma rhinocerontis* (Owen) was obtained by genome skimming from next-generation sequencing (NGS) data following Crampton-Platt *et al.* (2015).

BLAST search (Altschul *et al.*, 1990), MITOS search (Bernt *et al.*, 2013) and DNAMAN software (v8; Lynnon Corp., San Ramon, CA, U.S.A.) with another Oestroidea mitogenome as reference were used to identify genomic positions and gene boundaries of protein-coding genes (PCGs), ribosomal RNA (rRNA) and transfer RNA genes (tRNA). Nucleotide composition and codon usage were calculated using MEGA7 (Kumar *et al.*, 2016).

Base composition and substitution rates

Mitogenomes of 46 calyptrate species (File S3) were used for base composition estimation. Base compositions of each of the PCGs (for all three codon positions together and for each codon positions separately), rRNA and tRNA genes, concatenated PCGs, rRNA and tRNA genes, and complete or nearly complete mitogenomes were calculated using MEGA7 (Kumar *et al.*, 2016).

For calyptrate families with more than one mitogenome available, we selected one species from each genus (File S3) and calculated family-specific Tamura–Nei (TN) substitution rates using MEGA7. Pairwise distances of all 13 PCGs for substitutions at fourfold degenerate sites, and numbers of synonymous and nonsynonymous substitution per site were calculated following Eo & DeWoody (2010), and subsequently divided by divergence time, with divergence time of each family obtained from Cerretti *et al.* (2017).

Phylogenetic analysis

A total of nine oestrid mitogenomes, including the previously documented mitogenome of *Ga. pecorum* (Fabricius) (Zhang *et al.*, 2016a) plus mitogenomes from one exemplar species from each of the bot fly subfamilies Hypodermatinae [*Hypoderma lineatum* (Villers)] and Cuterebrinae [*Dermatobia hominis* (Linnaeus)], were included in the present study (Table 2).

Table 2. Taxon sampling in the present study.

Species	GenBank accession number	Reference
Dermatobia hominis	NC_006378	Azeredo-Espin *et al.* (2004)
Hypoderma lineatum	NC_013932	Weigl *et al.* (2010)
Gasterophilus pecorum	NC_029812	Zhang *et al.* (2016a)
Gasterophilus haemorrhoidalis	MG920502	Present study
Gasterophilus inermis	MG920503	Present study
Gasterophilus intestinalis	MG920504	Present study
Gasterophilus nasalis	MG920505	Present study
Gasterophilus nigricornis	MG920506	Present study
Gyrostigma rhinocerontis	MK045312	Present study
Cobboldia loxodontis	MK045310–MK045311	Present study

Phylogenetic analyses were conducted using the 13 PCGs and two rRNA genes. Each mitochondrial gene was aligned separately using MAFFT v.7.310 (Katoh & Standley, 2013). For PCGs, the option L-INS-i was used, with the iterative refinement method incorporating local pairwise alignment information (--localpair). After aligning, all alignments were translated into amino acid sequences in MEGA7 and adjusted to ensure reading frame fidelity. The same aligning parameters were used for rRNA genes, except that Q-INS-i was used, as the secondary structure of RNA is considered by this strategy. Individual alignments were then concatenated into a final matrix using SEQUENCEMATRIX v.1.8 (Vaidya *et al.*, 2011).

Phylogenetic trees were generated using Bayesian inference (BI) and maximum likelihood (ML), with dataset partitioned by gene. The best partitioning scheme and substitution model for BI was evaluated using PARTITIONFINDER2 (Lanfear *et al.*, 2017), after the 'greedy' algorithm with branch lengths estimated as 'linked', following the corrected Akaike information criterion. Bayesian inference was then conducted at the CIPRES webserver (Miller *et al.*, 2010) (https://www.phylo.org/) using MRBAYES 3.2.6 (Ronquist & Huelsenbeck, 2003). Two independent runs were conducted, each with four chains (one cold and three hot chains), for 10 million generations, and samples were drawn every 1000 generations. The first 25% of steps were discarded as burn-in.

The ML analyses were performed using IQTREE (Nguyen *et al.*, 2015), based on the best partitioning strategy searched by the self-implemented MODELFINDER (Kalyaanamoorthy *et al.*, 2017). Node support values were estimated with standard bootstrap resampling. The resulting trees were visualized using the iTOL online tool (https://itol.embl.de; Letunic & Bork, 2016).

Table 3. Characters used for ancestral state estimation in the present study.

Species	Geographic distribution	Oviposition site	Third-instar larvae attachment location
Dermatobia hominis	Neotropic	Environment	Subdermal tissue
Hypoderma lineatum	Palaearctic	Nonhead area of host body	Subdermal tissue
Cobboldia loxodontis	Afrotropic/Afrotropic + Palaearctic + Oriental	Head area of host body	Stomach
Gyrostigma rhinocerontis	Afrotropic	Nonhead and head area of host body	Stomach
Gasterophilus pecorum	Afrotropic + Palaearctic	Environment	Pharynx + stomach
Gasterophilus haemorrhoidalis	Afrotropic + Palaearctic	Head area of host body	Large intestine
Gasterophilus inermis	Afrotropic + Palaearctic	Head area of host body	Large intestine
Gasterophilus ternicinctus	Afrotropic	Unknown	Stomach
Gasterophilus intestinalis	Palaearctic	Nonhead area of host body	Stomach
Gasterophilus nasalis	Afrotropic + Palaearctic	Head area of host body	Pylorus and duodenum
Gasterophilus meridionalis	Afrotropic	Unknown	Pylorus and duodenum
Gasterophilus nigricornis	Palaearctic	Head area of host body	Pylorus and duodenum

Reconstruction of ancestral states and ancestral distribution

As phylogenetically close outgroups and relatively dense sampling are crucial to determine the ancestral node states, especially for maximum parsimony (MP) reconstruction (Salisbury & Kim, 2001), *Gasterophilus ternicinctus* Gedoelst and *Gasterophilus meridionalis* (Piller & Evans) were added to the phylogeny based on the sparse existing morphological evidence in order to perform reconstructions on a complete taxon coverage (see Results).

Data on distribution and biology (Table 3; summarized in File S4) were collected from Zumpt (1965) and other relevant literature (e.g. Horak *et al.*, 1984). Parasites of domestic hosts will often have greatly expanded their geographic distribution along with that of their hosts, and we have attempted to code what we consider as original (i.e. pre-human) distributions. The distribution is accordingly coded as Palaearctic for *Ga. intestinalis* (De Geer) as well as for the outgroup *H. lineatum*.

STATISTICAL DISPERSAL-VICARIANCE ANALYSIS [S-DIVA (Yu *et al.*, 2010); modified from DIVA (Ronquist, 1997)] is often used for biogeographical reconstructions, based on the assumption that speciation is caused by vicariance and minimizing implied dispersal and extinction events (Ronquist, 1997). Such vicariance (or separation/isolation) during speciation could happen not only spatially, but also temporally (e.g. Filchak *et al.*, 2000), or along with habitat divergence (e.g. Linn *et al.*, 2003). Two developmental strategies of *Gasterophilus* spp. (choice of oviposition site and LIII attachment site) involve the physical position of specific immature stages, and, like geographical data, they can be considered as evolving through processes equivalent to dispersal and vicariance. S-DIVA was used for reconstructing ancestral distributions and the ancestral states for oviposition site and third-instar larval attachment site.

No phylogenetic hypothesis has been proposed previously for the four known species of *Cobboldia* [*C. elephantis* (Cobbold) (Oriental), *C. loxodontis* Brauer (Afrotropical), *C. roverei* Gedoelst (Afrotropical), *C. russanovi* (Palaearctic; extinct)]. This prevents a specific reconstruction of the ancestral area, i.e. of the distribution of the hypothetical ancestor of *Cobboldia*, and alternative analyses were therefore performed with the distribution of *Cobboldia* coded as either Afrotropical + Palaearctic + Oriental or Afrotropical.

Oviposition sites were divided into three different states: nonhost environment, nonhead area of host, and head area of host (Table 3). The distinction between head and nonhead area of a host assumes that eggs deposited on the head are closer to the mouth and thus the newly hatched larva have a higher probability of reaching the host mouth and alimentary canal. Ancestral states of oviposition sites were reconstructed using MP implemented in MESQUITE v.3.2, and Bayesian binary Markov chain Monte Carlo (BBM) modified from MRBAYES 3.1.2 (Ronquist & Huelsenbeck, 2003) performed in RASP (Yu *et al.*, 2015). The BBM methods followed the Jukes and Cantor (JC) model. Reconstructions using S-DIVA were not possible for the complete phylogeny, as this method requires a complete character coding (i.e. no missing states), and thus our analysis was performed without the two species whose oviposition sites are unknown.

The LIII attachment sites were divided into five states (subdermal tissue, pharynx, stomach, pylorus and duodenum, and large intestine), and reconstructed using MP and BBM as described earlier, with additional reconstruction in S-DIVA performed in RASP.

Results

General features of Gasterophilinae mitogenomes

The total length of the mitogenomes of five *Gasterophilus* species ranges from 14 590 to 14 854 bp (File S5). After assembling and annotating, they were registered in the GenBank database (assigned accession numbers are given in Table 2). Each mitogenome contains the usual 13 PCGs, 22 tRNA genes, two rRNA genes and a noncoding region [the trnI gene was not sequenced for *Ga. nigricornis* (Loew) due to technical difficulties]. Similar to other oestroid flies (e.g. Zhang *et al.*, 2016a; Yan *et al.*, 2017), most mitochondrial genes in the present study are encoded on the majority strand (J-strand) with 23 genes (nine PCGs and 14 tRNA genes), and 15 genes (five

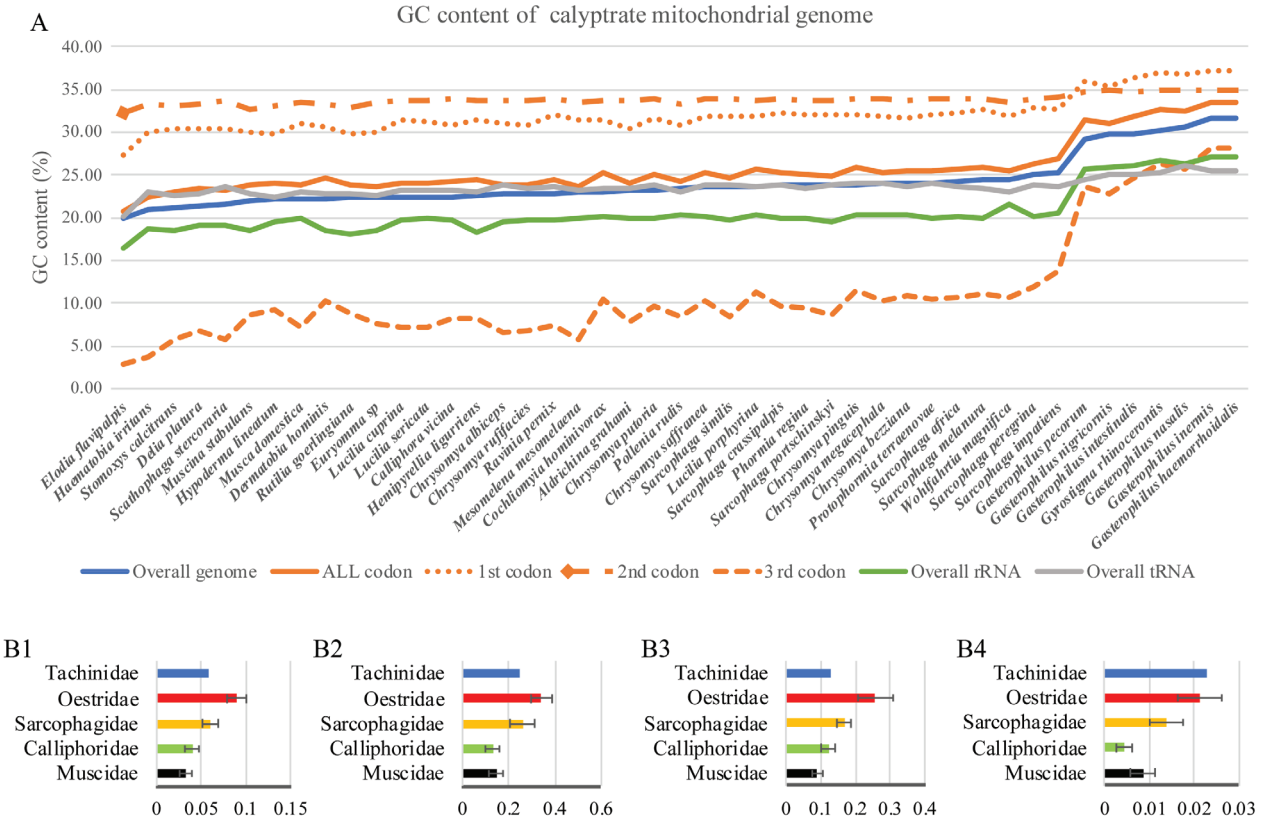

Fig. 1. Base composition and family-specific substitution rate of 13 mitochondrial protein-coding genes in calyptrates. (A) Guanine and cytosine (GC) content of calyptrate mitogenomes; (B1) mean evolutionary rates of each family estimated by dividing total amount of the Tamura–Nei substitution by divergence times; (B2) mean evolutionary rates of each family estimated by dividing total amount of substitution at fourfold degenerate sites by divergence times; (B3) mean evolutionary rates of each family estimated by dividing total amount of synonymous substitution by divergence times; (B4) mean evolutionary rates of each family estimated by dividing total amount of nonsynonymous substitution rates into divergence times. Error bars represent standard deviation from data of multiple species pairs. [Colour figure can be viewed at wileyonlinelibrary.com].

PCGs, eight tRNA genes and two rRNA genes) on the minority strand (N-strand) (File S5). The mitogenome of *Gy. rhinocerontis* obtained by genome skimming and mitochondrial genes of *C. loxodontis* are registered in GenBank (see Table 2 for accession numbers). Nearly all PCGs of *Gasterophilus* have one of the common start codons – ATG, TCG, ATA, or ATT – except for ATP8, which begins with ATC in *Ga. haemorrhoidalis* (Linnaeus) and *Ga. inermis* (Brauer), and with ATT in other *Gasterophilus* species. The majority of the PCGs terminate with TAA, TAG or T as stop codon, and only two PCGs (ND3 and ND4) have a slightly different termination (File S5).

Base composition and substitution rates

At the mitogenomic level, Gasterophilinae exhibit higher guanine and cytosine (GC) content than other oestrids and apparently all other calyptrates, and GC enrichment mainly occurs at third-codon positions of PCGs and rRNA genes (Fig. 1A). Mitogenomic GC content of calyptrates included in the present study ranges from 20.03% (*Elodia flavipalpis*) to 31.76% (*Ga. haemorrhoidalis*). The GC content of PCGs varies from 20.86% (*E. flavipalpis*) to 33.44% (*Ga. haemorrhoidalis*), and that of the first-, second- and third-codon positions ranges from 27.43% (*E. flavipalpis*) to 37.24% (*Ga. haemorrhoidalis*), 32.22% (*E. flavipalpis*) to 35.03% (*Ga. nasalis*), and 2.92% (*E. flavipalpis*) to 28.19% (*Ga. haemorrhoidalis*), respectively. *Elodia flavipalpis* has the lowest GC content of rRNA genes (16.51%) and tRNA genes (20.23%) whereas *Ga. haemorrhoidalis* and *Ga. inermis* have the highest GC content of rRNA genes (27.07%), and *Ga. nasalis* (Linnaeus) has the highest GC content of tRNA genes (26.01%).

Oestridae (0.09×10^{-10}) show a much higher TN substitution rate than other calyptrate families (Fig. 1B1–B4). Substitution rates at fourfold degenerate sites and synonymous sites show the same pattern as the TN rate, with the highest rate occurring in Oestridae, whereas for substitution rates at nonsynonymous sites, Tachinidae show the highest evolutionary rate, followed by Oestridae, Sarcophagidae, Muscidae and Calliphoridae.

Phylogeny of Gasterophilinae

The ML and BI (average standard deviation of split frequencies = 0.000 768; estimated sample size of all parameters

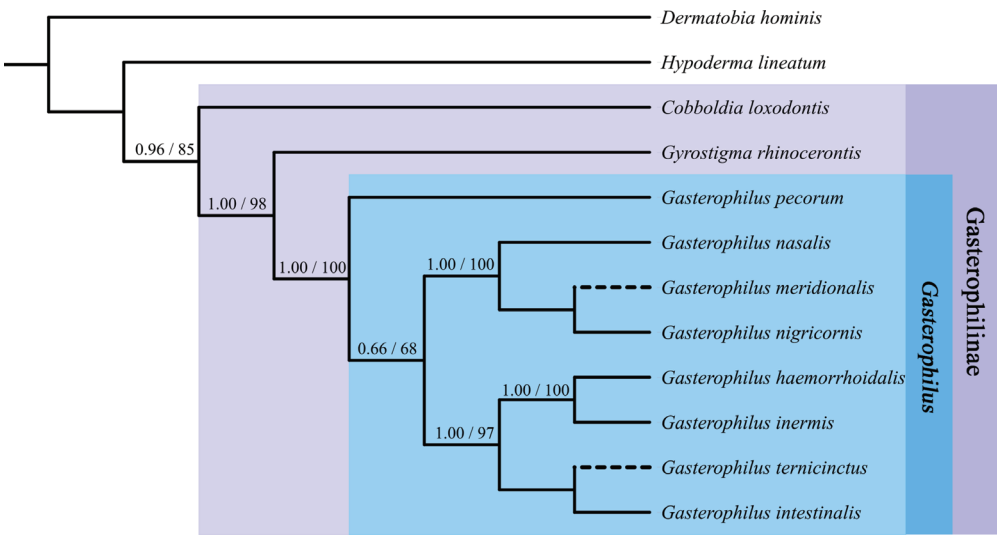

Fig. 2. Phylogeny of stomach bot flies inferred from mitogenomic data. Dashed branches indicate taxa placement based on morphological characters only. Numbers at the nodes are posterior probabilities (Bayesian trees)/bootstrap values (maximum likelihood trees) for molecular phylogeny construction. [Colour figure can be viewed at wileyonlinelibrary.com].

> 1000; potential scale reduction factor of all parameters = 1.00) analyses provide identical results, with Gasterophilinae being monophyletic, and a well-supported, monophyletic *Gasterophilus* being sister to the species of the rhino stomach bot fly, *Gy. rhinocerontis*. Within *Gasterophilus*, *Ga. pecorum* is estimated to be the sister group of the remaining horse stomach bot flies, although with only modest support (Fig. 2). The remaining *Gasterophilus* are split into two clades [*Ga. intestinalis* (*Ga. haemorrhoidalis* + *Ga. inermis*)] and (*Ga. nasalis* + *Ga. nigricornis*). The two Afrotropical endemics were placed on the cladogram based on their respective morphology. *Gasterophilus ternicinctus* was placed as sister to *Ga. intestinalis* based on their shared unique process on the hind trochanter (spatula-shaped in male and tubercular in female; Zumpt, 1965). *Gasterophilus meridionalis* was placed as sister to *Ga. nigricornis* based on shared features such as the first thoracic segment of LIII extended in a shelf-like manner over the pseudocephalon in *Ga. meridionalis*, *Ga. nasalis* and *Ga. nigricornis*; and *Ga. meridionalis* and *Ga. nigricornis* with a rugose base of the mouthhook (Zumpt, 1965; Colwell *et al.*, 2007; Li *et al.*, 2018)).

Ancestral area reconstruction

Under the present taxon sampling, the hypothetical ancestral distribution of the stomach bot flies cannot be more precise than the distribution observed for *Cobboldia* (e.g. Afrotropical + Oriental + Palaearctic) (Fig. 3; File S6). However, the Afrotropical region is reconstructed as the ancestral distribution both for the clade (*Gyrostigma* + *Gasterophilus*) and for all lineages of the *Gasterophilus* clade, except those leading to the clades (*Ga. intestinalis* + *Ga. ternicinctus*) and (*Ga. nigricornis* + *Ga. meridionalis*), for which the ancestral distribution in both cases are estimated as either Afrotropical or Palaearctic.

Estimates of the ancestral oviposition sites are similar between MP and BBM analyses (Fig. 4; File S6). The ancestral stomach bot fly most probably oviposited on the head of its host (62.54% probability in BBM estimation, and the only possibility in MP estimation). The ancestor of (*Gyrostigma* + *Gasterophilus*) probably oviposited on the head of the host according to the MP estimation, while the BBM estimation favoured either the combined head + nonhead area (45.66%) or only the head area (42.14%), with a 12.20% probability for an ancestral oviposition on a nonhost (i.e. environmental) substrate. The ancestral horse stomach bot fly is also estimated to have most probably oviposited on the head area of the host, which is also the estimate obtained for most of the subordinate *Gasterophilus* nodes. By contrast, the S-DIVA estimation indicated the ancestral oviposition site of the ancestor of the clade [*Ga. intestinalis* (*Ga. haemorrhoidalis* + *Ga. inermis*)] to be the head + nonhead area of the host.

The most likely ancestral LIII attachment site for all stomach bot flies, for (*Gyrostigma* + *Gasterophilus*) and for *Gasterophilus*, is estimated to be the stomach, either as the only estimate or, for *Gasterophilus*, with a probability of 67.90%, which is much higher than the next highest probability of 27.81% for a position in the combined area of the pharynx and stomach (Fig. 5; File S6). The LIII attachment site of the ancestor of all *Gasterophilus* except for *Ga. pecorum* is estimated to most likely be the stomach in MP and BBM reconstructions, while it is either the stomach + pylorus-duodenum or the stomach + pylorus-duodenum + large intestine in S-DIVA reconstruction (node 18). Similar to the previous node, the estimation for the ancestor of [(*Ga. intestinalis* + *Ga. ternicinctus*), (*Ga. haemorrhoidalis* + *Ga. inermis*)] in S-DIVA (node 15, stomach + large intestine) is different from that in the MP and BBM

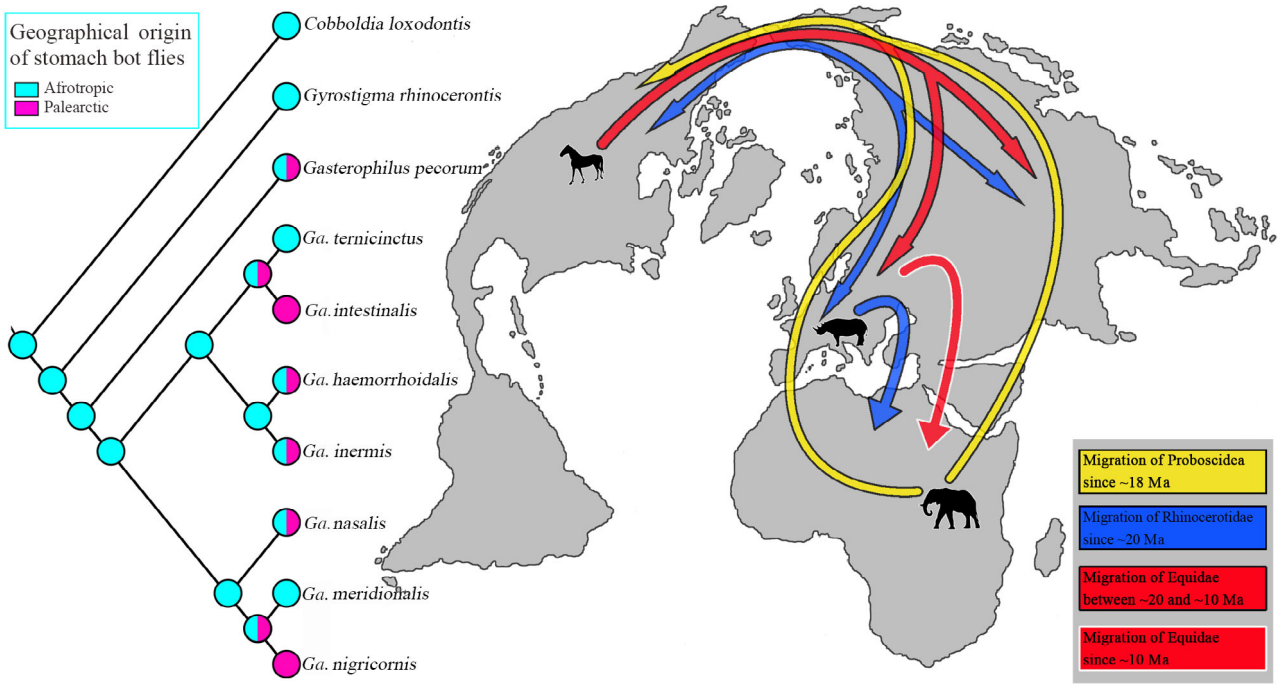

Fig. 3. Estimation of geographical origin of stomach bot flies. Left: STATISTICAL DISPERSAL-VICARIANCE ANALYSIS (S-DIVA) estimation based on proposed *Gasterophilus* species tree; right: migrating routes of Proboscidea, Rhinocerotidae and Equidae since 20 Ma, as indicated by the literature (Bernor *et al.*, 2010; Geraads, 2010; Sanders *et al.*, 2010; Franzen & Brown, 2011). [Colour figure can be viewed at wileyonlinelibrary.com].

estimations (most likely is the stomach). Estimations for the remaining nodes are identical.

Discussion

We reconstructed the stomach bot fly phylogeny using the most extensive molecular data and broadest taxon coverage currently available, including complete mitochondrial genomes from six of the known eight species of *Gasterophilus*, and for the first time with mitogenomic data for *Gyrostigma* and *Cobboldia*. The present phylogeny is similar to the morphology-based phylogenetic topology of Zhang *et al.* (2016b) in recovering *Ga. pecorum* as sister to the remaining *Gasterophilus*, and a well-supported sister pairing of *Ga. nasalis* and *Ga. nigricornis*. However, Zhang *et al.* (2016b) differs by having *Ga. haemorrhoidalis* grouping with (*Ga. nasalis* + *Ga. nigricornis*) rather than with *Ga. intestinalis* in a clade with *Ga. inermis* and *Ga. ternicinctus* of the present study. The study by Zhang *et al.* (2016b) was based on a limited sample of 17 morphological characters with a narrow focus on the adult antenna, whereas the present study is based on complete mitogenomic data.

Host shifts

The evolution of host shifts is a significant issue in evolutionary biology (De Vienne *et al.*, 2013), and the conditions facilitating host shifts in parasites have attracted much attention (De Fine Licht, 2018). Assuming proboscideans were the ancestral hosts of early stomach bot flies (Pape, 2006), then a host shift probably occurred from elephants to rhinoceroses no earlier than *c.* 20 Ma, and then from rhinoceroses to equines no earlier than *c.* 10.5 Ma. Our study supports the idea that the ancestral distribution of stomach bot flies is the Afrotropical region (Pape, 2006). The arrival of the earliest Afrotropical rhinoceroses (Rhinocerotidae: *Brachypotherium*) (Geraads, 2010) from the northern continents *c.* 20 Ma provided opportunities for ancient elephant stomach bot flies to diversify by colonizing a rhinoceros host. Equids were absent from the Afrotropics until *c.* 10.5 Ma (Bernor *et al.*, 2010; Franzen & Brown, 2011), and host shifts of stomach bot flies from rhinos to equids must have occurred thereafter. These host shifts may have been facilitated by the habitat preferences and feeding habits of the hosts and thus the likelihood of oviposition mistakes. In the Afrotropics, Miocene elephants were mixed feeders with a dominant preference for browsing until *c.* 7.5 Ma (Cerling *et al.*, 1999), whereas rhinoceroses recorded from Africa (Geraads, 2010) were mainly grazers (Prothero *et al.*, 1989), and equids had a dominant preference for grazing (Prothero *et al.*, 1989; Bernor *et al.*, 2010). This means that horses arriving in the Afrotropics were more likely to be in the vicinity of rhinoceroses than elephants, and thus the more likely target for oviposition mistakes. After colonizing new equid hosts, the diversification of the stomach bot flies led to the appearance of present-day species of *Gasterophilus*.

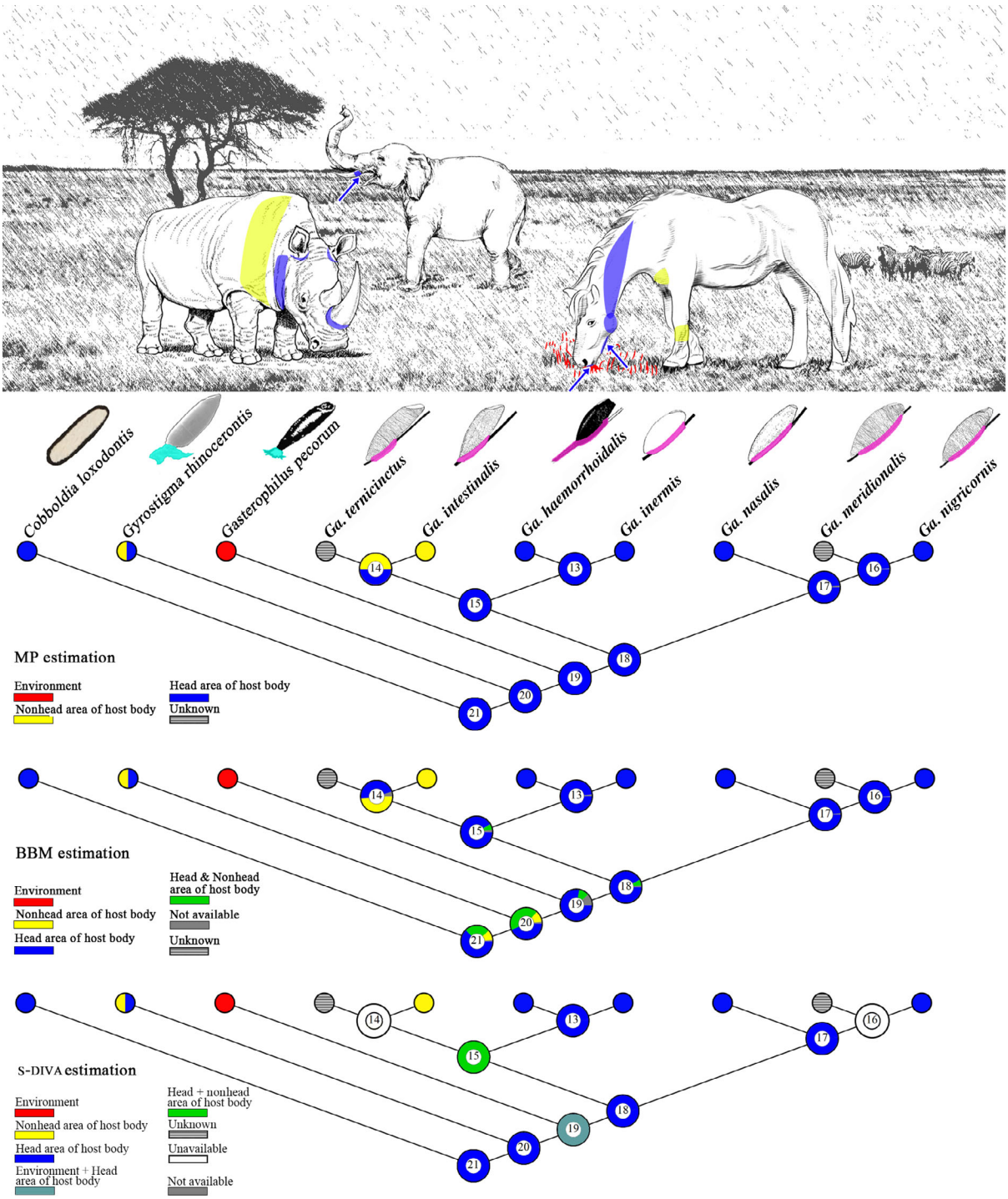

Fig. 4. Reconstructions of ancestral oviposition sites of stomach bot flies. Cladograms from top to bottom are reconstruction using maximum parsimony (MP), Bayesian binary Markov chain Monte Carlo (BBM) and STATISTICAL DISPERSAL-VICARIANCE ANALYSIS (S-DIVA), respectively. Possibility (MP) or probability of each possibility (BBM and S-DIVA) for each node is shown as a colour proportion on the node. Eggs above cladograms are *Cobboldia elephantis* [this is the only egg of *Cobboldia* recorded in the literature (Patton, 1922)], *Gyrostigma rhinocerontis*, *Gasterophilus pecorum*, *Gasterophilus ternicinctus*, *Gasterophilus intestinalis*, *Gasterophilus haemorrhoidalis*, *Gasterophilus inermis*, *Gasterophilus nasalis*, *Gasterophilus meridionalis* and *Gasterophilus nigricornis*, from left to right, with the structure of a type II attachment organ (AO) highlighted in Arctic blue, and that of a type I AO in pink. Known oviposition sites of each species are illustrated in the top picture using different colours. [Colour figure can be viewed at wileyonlinelibrary.com].

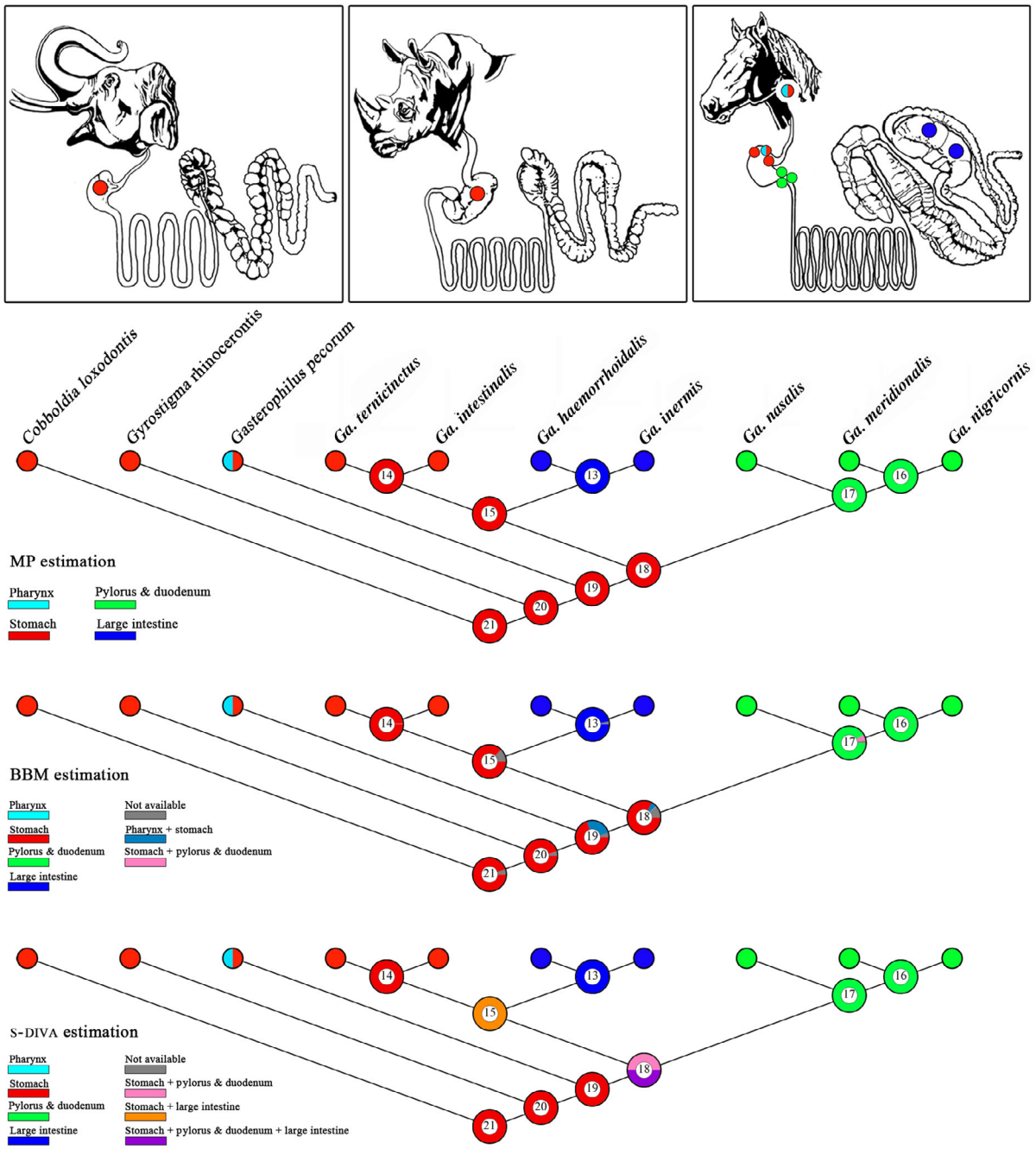

Fig. 5. Reconstructions of ancestral third instar larvae attaching locations of stomach bot flies. Cladograms from top to bottom are reconstructed using maximum parsimony (MP), Bayesian binary Markov chain Monte Carlo (BBM) and STATISTICAL DISPERSAL-VICARIANCE ANALYSIS (S-DIVA), respectively. Possibility (MP) or probability of each possibility (BBM and S-DIVA) for each node is shown as a colour proportion on the node. Digestive tracts of mammals are modified from Stevens & Hume (1998). [Colour figure can be viewed at wileyonlinelibrary.com].

Origin and dispersal

Our study suggests that stomach bot flies adapted to changes in the habitat imposed by the abiotic environment, rather than in response to migrating with their hosts. An Afrotropical origin of *Gasterophilus* is surprising because equids entered Africa as late as *c*. 10.5 Ma (Janis, 1993; Bernor *et al.*, 2010; Franzen & Brown, 2011), and ancient elephants or rhinoceroses infected with stomach bot flies would have had ample opportunities to associate with horses that were very common in Eurasia

during the period 18–10.5 Ma (Franzen & Brown, 2011). One explanation is that ancient stomach bot flies were strictly confined to tropical habitats, which would fit well with the notion that early bot fly evolution most probably took place in a humid tropical forest environment (Pape, 2006). Thus, the early stomach bot flies may not have entered Eurasia with their hosts because the environment was not suitable for the adult flies. Nevertheless, the ancient lineages of stomach bot flies must have adapted to the changing environment in the Afrotropical region, which became drier and cooler with the spread of savannahs since the mid-Miocene (Feakins & Demenocal, 2010), and the ancestral *Gasterophilus* would have succeeded in colonizing early equids that had migrated into this region. The presence of *Cobboldia russanovi* in a woolly mammoth preserved in the permafrost of the Siberian tundra (Grunin, 1973) might be evidence that this adaptive shift also occurred in the genus *Cobboldia* and is consistent with the view that stomach bot flies adapted in response to temporal changes in the abiotic environments rather than to spatial changes imposed with their early migrating hosts. The presence of *Gyrostigma sumatrensis* Brauer in the Sumatran rhinoceros [*Dicerorhinus sumatrensis* (Fischer)] may also fit this pattern. Although rhinoceroses used to be distributed in both Africa and Eurasia (Kalb *et al.*, 1982; Benefit & Monte, 1989), these Eurasian species may have been clear of stomach bot flies until a lineage evolved the capacity to survive as adults in a drier environment and subsequently spread into Eurasia [cf. fossil records: Kaya *et al.* (2012); Kazanci *et al.* (1999); Lehmann (1984)], eventually colonizing the Sumatran rhinoceros before African and Eurasian rhinoceroses were completely isolated from each other (Prothero, 1993; Pandolfi & Tagliacozzo, 2015).

Adaptations for oviposition

Our phylogeny identifies the evolution of several adaptations associated with changes in host, and in oviposition and larval attachment sites. The head area was the primary oviposition site for all stomach bot fly lineages, and the morphology of the attachment organ on the eggs of stomach bot flies appears to be linked to their specific oviposition locations. The variation in fecundity (Table 1) of *Gasterophilus* is consistent with the balanced mortality life-history strategy associated with the mode of infection of LI (Stearns, 1992; Mayhew, 2016). The larvae of stomach bot fly species that oviposit on the head must reach the mouth of the host by active migration and these adults lay fewer eggs than species that oviposit away from the head, where larvae can only enter the host body by being ingested. Interestingly, the fecundity of *Gy. rhinocerontis* is similar to that recorded for *Ga. intestinalis* – both produce about 700 eggs (Rodhain, 1915; Zumpt, 1965) – and both have similar oviposition sites on their respective rhinoceros and equine hosts.

Evolution in attachment site of third instar larvae

According to our reconstruction, the stomach is likely to have been the preferred LIII attachment site of early stomach bot fly lineages. Our phylogeny reveals two shifts in the main location of LIII during the diversification of *Gasterophilus*: once from the stomach and into the large intestine, and once from the stomach and into the pylorus and duodenum. Attachment site selection by parasites is an active process, potentially leading to niche specialization, and a change in feeding strategy is thought to be one of the primary causes of differentiation in attachment site selection (Petter, 1962; Schad, 1963; Holmes, 1972). Similarly, the oral plates (Li *et al.*, 2018), i.e. the cephaloskeletal organ in *Gasterophilus* LIII that assists in feeding, varies (Principato, 1986), although species with the same attachment site have similar oral plate morphology. It is noteworthy that LIII of the species attaching to the pylorus and duodenum only have a single row of spines on each body segment (if there are spines at all), whereas other species of *Gasterophilus* have two to three rows of spines (Li *et al.*, 2018). The configuration and arrangement of body spines may be related to the attachment site, considering that peristalsis varies between stomach, duodenum and the large intestine (Van Weyenberg *et al.*, 2006; Huizinga & Lammers, 2008).

Molecular adaptation of stomach bot flies

Our study suggests that an intestinal parasitoid lifestyle has contributed to the evolution of genomic base composition in the stomach bot flies. The species of Gasterophilinae included in the present study have much higher GC content (in terms of overall mitogenome, PCGs and rRNA genes) than other calyptrates. Although the cause of GC content variation among and within genomes of organisms is still unclear, an environmental influence on the nucleotide composition of microbial genomes has been documented (Foerstner *et al.*, 2005). The thermal adaptation hypothesis (Bernardi, 1995) argues that G:C pairs, which are connected by three hydrogen bonds, are more thermally stable than A:T pairs, which are connected by only two hydrogen bonds (Wada & Suyama, 1986), and GC-rich genomes should accordingly ensure a more reliable protein synthesis due to more stable mRNA transcripts under higher temperatures (Bernardi, 1995). Consequently, high GC content could be expected to be found more often in animals with high body temperature (Bernardi, 2000; Mooers & Holmes, 2000). Although the thermal adaptation hypothesis has been questioned (e.g. Vinogradov & Anatskaya, 2017), our study lends some support to this hypothesis, especially as the accumulation of GC mainly happens at freely evolving sites (i.e. third-codon position of PCGs) and rRNA genes (Hurst & Merchant, 2001).

It is also believed that GC content covaries with mutation rate (Vinogradov & Anatskaya, 2017; Kiktev *et al.*, 2018), and that GC at third-codon positions therefore accumulates in fast-evolving lineages (Romiguier *et al.*, 2010). Therefore, the enrichment of GC in overall mitogenomes and at third-codon positions of PCGs of stomach bot flies indicates their higher evolutionary rate compared with other calyptrates. This is possibly an effect of the evolutionary arms race between stomach bot flies and their hosts as implied by the Red Queen hypothesis (Van Valen, 1973) and as a means to survive in the challenging environment of the mammalian digestive tract.

Conclusions

Stomach bot flies, the larvae of which are obligate gastrointestinal parasitoids of mammals, represent a very specialized group of animals, whose evolutionary history is only recently being illuminated through modern molecular/phylogenomic methodologies. The unusual lifestyle of these flies, several species of which face a high risk of extinction (Colwell *et al.*, 2009), can provide significant insights into the evolution of parasitism in insects. We use mitogenomics to reveal an evolutionary history of the horse stomach bot flies, *Gasterophilus* spp. We estimate that all stomach bot fly lineages (elephant, rhino, and horse stomach bot flies) originated in the Afrotropics, with host shifts from elephants to rhinos and from rhinos to horses, and a subsequent dispersal into the Palaearctic and Oriental regions with their hosts. The head area is the ancestral oviposition site for all stomach bot fly lineages, and changes in fecundity evolved according to the probability of the first-instar larvae to enter the host mouth. The stomach is the ancestral attachment site for third-instar larvae of all stomach bot fly lineages. A high GC content may represent a molecular adaptation to life in the high-temperature or acid environment of the mammalian digestive tract.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

File S1. Specimens used for DNA extraction in the present study.

File S2. Primer pairs (5′-3′), size of amplicons and protocols used for amplifying mitogenomes.

File S3. Taxon sampling for calculating guanine and cytosine (GC) content of calyptrate mitogenomes.

File S4. Biology information of stomach bot flies collected from literature.

File S5. Organization of the mitogenome of Gasterophilinae in the present study.

File S6. Matrices and results of ancestral states estimation.

Acknowledgements

We are grateful to Prof. Rudolf Meier (National University of Singapore, Singapore), Prof. Thierry Bourgoin (Muséum National d'Histoire Naturelle, France), Prof. Adrian Lister (Natural History Museum, U.K.) and Dr. Juha Saarinen (University of Helsinki, Finland) for kindly providing valuable comments to the manuscript. This study was supported by the Fundamental Research Funds for the Central Universities (no. JC2015-04) and the National Natural Science Foundation of China (nos 31572305 and 31872964). The authors declare no conflict of interest.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Azeredo-Espin, A.M.L., Junqueira, A.C.M., Lessinger, A.C., Lyra, M.L., Torres, T.T., da Silva, F.R., Meirelles, F.V & Maia, A.A.M. (2004) The complete mitochondrial genome of the human botfly *Dermatobia hominis* (Diptera: Oestridae). *Unpublished Poster, ESA Annual Meeting and Exhibition*. Available at: https://Esa.Confex.Com/Esa/2004/Techprogram/Paper_16801.htm [accessed on 9 March 2019].

Balashov, Y.S. (2006) Types of parasitism of acarines and insects on terrestrial vertebrates. *Entomological Review*, **86**, 957–971.

Barker, F.K. (2014) Mitogenomic data resolve basal relationships among passeriform and passeridan birds. *Molecular Phylogenetics and Evolution*, **79**, 313–324.

Benefit, B.R. & Monte, L.M. (1989) New primate fossils from the middle Miocene of Maboko Island, Kenya. *Journal of Human Evolution*, **18**, 493–497.

Bernardi, G. (1995) The human genome: organization and evolutionary history. *Annual Review of Genetics*, **29**, 445–476.

Bernardi, G. (2000) Isochores and the evolutionary genomics of vertebrates. *Gene*, **241**, 3–17.

Bernor, R.L., Armour-Chelu, M.J., Gilbert, H., Karser, T.M. & Schulz, E. (2010) Equidae. *Cenozoic Mammals of Africa* (ed. by L. Werdelin and W.J. Sanders), pp. 685–721. University of California Press, Berkeley.

Bernt, M., Donath, A., Jühling, F. *et al.* (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, **69**, 313–319.

Cameron, S.L. (2014) Insect mitochondrial genomics: implications for evolution and phylogeny. *Annual Review of Entomology*, **59**, 95–117.

Cerling, T.E., Harris, J.M. & Leakey, M.G. (1999) Browsing and grazing in elephants: the isotope record of modern and fossil proboscideans. *Oecologia*, **120**, 364–374.

Cerretti, P., Stireman, J.O., Pape, T., O'Hara, J.E., Marinho, M.A.T., Rognes, K. & Grimaldi, D.A. (2017) First fossil of an oestroid fly (Diptera: Calyptratae: Oestroidea) and the dating of oestroid divergences. *PLoS One*, **12**, e0182101.

Colwell, D. (2006) Life cycle strategies. *The Oestrid Flies: Biology, Host-Parasite Relationships, Impact and Management* (ed. by D.D. Colwell, M.J. Hall and P.J. Scholl), pp. 67–77. CABI, Oxfordshire.

Colwell, D.D., Otranto, D. & Horak, I.G. (2007) Comparative scanning electron microscopy of *Gasterophilus* third instars. *Medical and Veterinary Entomology*, **21**, 255–264.

Colwell, D.D., Otranto, D. & Stevens, J.R. (2009) Oestrid flies: eradication and extinction versus biodiversity. *Trends in Parasitology*, **25**, 500–504.

Courtney, G.W., Pape, T., Skevington, J.H. & Sinclair, B.J. (2017) Biodiversity of Diptera. *Insect Biodiversity* (ed. by R.G. Foottit and P.H. Adler), pp. 229–278. Wiley-Blackwell, Oxford.

Crampton-Platt, A., Timmermans, M.J.T.N., Gimmel, M.L., Kutty, S.N., Cockerill, T.D., Khen, C.V. & Vogler, A.P. (2015) Soup to tree: the phylogeny of beetles inferred by mitochondrial metagenomics of a Bornean rainforest sample. *Molecular Biology and Evolution*, **32**, 2302–2316.

De Fine Licht, H.H. (2018) Does pathogen plasticity facilitate host shifts? *PLoS Pathogens*, **14**, e1006961.

De Meeûs, T. & Renaud, F. (2002) Parasites within the new phylogeny of eukaryotes. *Trends in Parasitology*, **18**, 247–250.

De Vienne, D.M., Refrégier, G., López-Villavicencio, M., Tellier, A., Hood, M.E. & Giraud, T. (2013) Cospeciation vs host-shift speciation:

methods for testing, evidence from natural associations and relation to coevolution. *New Phytologist*, **198**, 347–385.

Eo, S.H. & DeWoody, J.A. (2010) Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3587–3592.

Faucci, A., Toonen, R.J. & Hadfield, M.G. (2007) Host shift and speciation in a coral-feeding nudibranch. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 111–119.

Feakins, S.J. & Demenocal, P.B. (2010) Global and African regional climate during the Cenozoic. *Cenozoic Mammals of Africa* (ed. by L. Werdelin and W.J. Sanders), pp. 45–55. University of California Press, Berkeley.

Feener, D.H. Jr & Brown, B.V. (1997) Diptera as parasitoids. *Annual Review of Entomology*, **42**, 73–97.

Filchak, K.E., Roethele, J.B. & Feder, J.L. (2000) Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature*, **407**, 739–742.

Foerstner, K.U., von Mering, C., Hooper, S.D. & Bork, P. (2005) Environments shape the nucleotide composition of genomes. *EMBO Reports*, **6**, 1208–1213.

Franzen, J.L. & Brown, M. (2011) *The Rise of Horses: 55 Million Years of Evolution*, pp. 1–211. The Johns Hopkins University press, Baltimore.

Geraads, D. (2010) Rhinocerotidae. *Cenozoic Mammals of Africa* (ed. by L. Werdelin and W.J. Sanders), pp. 669–683. University of California press, Berkeley.

Gowda, A.K.J., Dharanesha, N.K., Giridhar, P. & Byre Gowda, S.M. (2017) *Cobboldia elephantis* (Cobbold, 1866) larval infestation in an Indian elephant (*Elephas maximus*). *Journal of Parasitic Diseases*, **41**, 364–366.

Grimaldi, D. & Engel, M.S. (2005) *Evolution of the Insects*, pp. 1–732. Cambridge University Press, New York.

Grunin, K. (1973) First discovery of larvae of the mammoth bot-fly *Cobboldia* (*Mamontia*, subgen. n.) *russanovi*, sp. n. (Diptera, Gasterophilidae). *Entomological Review*, **52**, 165–169.

Guimarães, J.H. & Papavero, N. (1999) *Myiasis in Man and Animals in the Neotropical Region: Bibliographic Database*, pp. 1–308. Plêiade, São Paulo.

Holmes, J.C. (1972) Site selection by parasitic helminths: interspecific interactions, site segregation, and their importance to the development of helminth communities. *Canadian Journal of Zoology*, **51**, 333–347.

Horak, I.G., de Vos, V. & de Klerk, B.D. (1984) Parasites of domestic and wild animals in South Africa. XVII. Arthropod parasites of Burchell's zebra, *Equus burchelli*, in the eastern Transvaal Lowveld. *Onderstepoort Journal of Veterinary Research*, **51**, 145–154.

Horreo, J.L. (2017) Revisiting the mitogenomic phylogeny of Salmoninae: new insights thanks to recent sequencing advances. *PeerJ*, **5**, e3828.

Huizinga, J.D. & Lammers, W.J.E.P. (2008) Gut peristalsis is governed by a multitude of cooperating mechanisms. *AJP: Gastrointestinal and Liver Physiology*, **296**, G1–G8.

Hurst, L.D. & Merchant, A.R. (2001) High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 493–497.

Janis, C.M. (1993) Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics*, **24**, 467–500.

Kalb, J.E., Jolly, C.J., Mebrate, A. *et al.* (1982) Fossil mammals and artefacts from the middle Awash Valley, Ethiopia. *Nature*, **298**, 25–29.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A. & Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587–589.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Kaya, T.T., Mayda, S., Kostopoulos, D.S. *et al.* (2012) Şerefköy-2, a new late Miocene mammal locality from the Yataĝan formation, Muĝla, SW Turkey. *Comptes Rendus Palevol*, **11**, 5–12.

Kazanci, N., Sen, S., Seyitoglu, G. *et al.* (1999) Geology of a new late Miocene mammal locality in Central Anatolia, Turkey. *Comptes Rendus de l'Académie des Sciences-Series IIA-Earth and Planetary Science*, **329**, 503–510.

Kiktev, D.A., Sheng, Z., Lobachev, K.S. & Petes, T.D. (2018) GC content elevates mutation and recombination rates in the yeast *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences*, **115**, E7109–E7118.

Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**, 1870–1874.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2017) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**, 772–773.

Lehmann, U. (1984) Notiz über Säugetierreste von der Insel Samos in der Sammlung des Geologisch-Paläontologischen Instituts und Museums Hamburg. *Mitteilungen aus dem Geologischen-Paläontologischen Institut der Universität Hamburg*, **57**, 147–156.

Letunic, I. & Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, **44**(W1), W242–W245.

Li, X., Chen, Y., Wang, Q., Li, K., Pape, T. & Zhang, D. (2018) Molecular and morphological characterization of third instar Palaearctic horse stomach bot fly larvae (Oestridae: Gasterophilinae, *Gasterophilus*). *Veterinary Parasitology*, **262**, 56–74.

Linn, C., Feder, J.L., Nojima, S., Dambroski, H.R., Berlocher, S.H. & Roelofs, W. (2003) Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences*, **100**, 11490–11493.

Mayhew, P.J. (2016) Comparing parasitoid life histories. *Entomologia Experimentalis et Applicata*, **159**, 147–162.

Miller, M.A., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES science gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*, pp. 1–8. New Orleans, Louisiana, USA. Available at: https://ieeexplore.ieee.org/abstract/document/5676129 [accessed on 9 March 2019].

Mooers, A.Ø. & Holmes, E.C. (2000) The evolution of base composition and phylogenetic inference. *Trends in Ecology and Evolution*, **15**, 365–369.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.

Nufio, C.R. & Papaj, D.R. (2004) Superparasitism of larval hosts by the walnut fly, *Rhagoletis juglandis*, and its implications for female and offspring performance. *Oecologia*, **141**, 460–467.

Pandolfi, L. & Tagliacozzo, A. (2015) *Stephanorhinus hemitoechus* (Mammalia, Rhinocerotidae) from the late Pleistocene of Valle Radice (Sora, Central Italy) and re-evaluation of the morphometric variability of the species in Europe. *Geobios*, **48**, 169–191.

Pape, T. (2006) Phylogeny and evolution of the bot flies. *The Oestrid Flies: Biology, Host-Parasite Relationships, Impact and Management*

(ed. by D.D. Colwell, M.J. Hall and P.J. Scholl), pp. 20–50. CABI, Oxfordshire.

Pape, T., Blagoderov, V. & Mostovski, M.B. (2011) Order Diptera Linnaeus, 1758. Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness (ed. by Z.-Q. Zhang). *Zootaxa*, **3148**, 222–229.

Pape, T., Piwczyński, M., Wyborska, D., Akbarzadeh, K. & Szpila, K. (2017) A new genus and species of hypodermatine bot flies (Diptera: Oestridae). *Systematic Entomology*, **42**, 387–398.

Patton, M.W.S. (1922) Note on the egg-laying habits of *Cobboldia elephantis*, Cobbold. *Indian Journal of Medical Research*, **10**, 63–65.

Petter, A.J. (1962) Redescription et analyse critique de quelques espèces d'Oxyures de la tortue grecque (*Testudo graeca* L.). Diversité des structures céphaliques (II). *Annales de Parasitologie Humaine et Comparée*, **37**, 140–152.

Principato, M. (1986) A study on the maxillae of third instar larvae of the six Italian *Gasterophilus* species (Diptera: Gasterophilidae): observations on their taxonomic significance in the identification to the species. *Bulletin de la Société Française de Parasitologie*, **1**, 121–130.

Prothero, D.R. (1993) Fifty million years of rhinoceros evolution. *Rhinoceros biology and conservation: Proceedings of the International Rhino Conference* (ed. by O.A. Ryder), pp. 82–91. San Diego Zoological Society, California, San Diego.

Prothero, D.R., Guerin, C. & Manning, E. (1989) The history of the Rhinocerotoidea. *The Evolution of Perissodactyls* (ed. by D.R. Prothero and R.M. Schoch), pp. 321–340. Oxford university press, New York.

Read, C.P. (1972) *Animal Parasitism*, pp. 1–192. Prentice Hall, Englewood Cliffs.

Rodhain, J. (1915) Note on oviposition by Oestrids of the genera *Gyrostigma* and *Cobboldia*. *Bulletin de la Société de Pathologie Exotique*, **8**, 275–279.

Romiguier, J., Ranwez, V., Douzery, E.J.P. & Galtier, N. (2010) Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Research*, **20**, 1001–1009.

Ronquist, F. (1997) Dispersal-Vicariance analysis: a new approach to the quantification of historical biogeography. *Systematic Biology*, **46**, 195–203.

Ronquist, F. & Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Salisbury, B. & Kim, J. (2001) Ancestral state estimation and taxon sampling density. *Systematic Biology*, **50**, 557–564.

Sanders, W.J., Gheerbrant, E., Harris, J.M., Saegusa, H. & Delmer, C. (2010) Proboscidea. *Cenozoic Mammals of Africa* (ed. by L. Werdelin and W.J. Sanders), pp. 161–251. University of California press, Berkeley.

Schad, G.A. (1963) Niche diversification in a parasitic species flock. *Nature*, **198**, 404–406.

Stearns, S.C. (1992) *The Evolution of Life Histories*, pp. 1–249. Oxford University Press, New York.

Stevens, C.E. & Hume, I.D. (1998) Contributions of microbes in vertebrate gastrointestinal tract to production and conservation of nutrients. *Physiological Reviews*, **78**, 393–427.

Stireman, J.O. III, O'Hara, J.E. & Wood, D.M. (2006) Tachinidae: evolution, behavior, and ecology. *Annual Review of Entomology*, **51**, 525–555.

Stireman, J.O. III, Cerretti, P., O'Hara, J.E., Blaschke, J.D. & Moulton, J.K. (2019) Molecular phylogeny and evolution of world tachinidae (Diptera). *Molecular Phylogenetics and Evolution*. https://doi.org/10.1016/j.ympev.2018.12.002.

Timmermans, M.J.T.N., Lees, D.C. & Simonsen, T.J. (2014) Towards a mitogenomic phylogeny of Lepidoptera. *Molecular Phylogenetics and Evolution*, **79**, 169–178.

Vaidya, G., Lohman, D.J. & Meier, R. (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*, **27**, 171–180.

Van Valen, L. (1973) A new evolutionary theory. *Evolutionary Theory*, **1**, 1–30.

Van Weyenberg, S., Sales, J. & Janssens, G.P.J. (2006) Passage rate of digesta through the equine gastrointestinal tract: a review. *Livestock Science*, **99**, 3–12.

Vinogradov, A.E. & Anatskaya, O.V. (2017) DNA helix: the importance of being AT-rich. *Mammalian Genome*, **28**, 455–464.

Wada, A. & Suyama, A. (1986) Local stability of DNA and RNA secondary structure and its relation to biological functions. *Progress in Biophysics and Molecular Biology*, **47**, 113–157.

Weigl, S., Testini, G., Parisi, A., Dantas-Torres, F., Traversa, D., Colwell, D.D. & Otranto, D. (2010) The mitochondrial genome of the common cattle grub, *Hypoderma lineatum*. *Medical and Veterinary Entomology*, **24**, 329–335.

Weinstein, S.B. & Kuris, A.M. (2016) Independent origins of parasitism in Animalia. *Biology Letters*, **12**, 20160324.

Wiegmann, B.M., Trautwein, M.D., Winkler, I.S. *et al.* (2011) Episodic radiations in the fly tree of life. *Proceedings of the National Academy of Sciences*, **108**, 5690–5695.

Yan, L., Zhang, M., Gao, Y., Pape, T. & Zhang, D. (2017) First mitogenome for the subfamily Miltogramminae (Diptera: Sarcophagidae) and its phylogenetic implications. *European Journal of Entomology*, **114**, 422–429.

Yu, Y., Harris, A.J. & He, X. (2010) S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. *Molecular Phylogenetics and Evolution*, **56**, 848–850.

Yu, Y., Harris, A.J., Blair, C. & He, X. (2015) RASP (reconstruct ancestral state in phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution*, **87**, 46–49.

Zelmer, D.A. (1998) An evolutionary definition of parasitism. *International Journal for Parasitology*, **28**, 531–533.

Zhang, D., Yan, L., Zhang, M. *et al.* (2016a) Phylogenetic inference of calyptrates, with the first mitogenomes for Gasterophilinae (Diptera: Oestridae) and Paramacronychiinae (Diptera: Sarcophagidae). *International Journal of Biological Sciences*, **12**, 489–504.

Zhang, D., Li, X., Liu, X., Wang, Q. & Pape, T. (2016b) The antenna of horse stomach bot flies: morphology and phylogenetic implications (Oestridae, Gasterophilinae: *Gasterophilus* Leach). *Scientific Reports*, **6**, 34409.

Zietara, M.S. & Lumme, J. (2002) Speciation by host switch and adaptive radiation in a fish parasite genus *Gyrodactylus* (Monogenea, Gyrodactylidae). *Evolution*, **56**, 2445–2458.

Zumpt, F. (1965) *Myiasis in Man and Animals in the Old World: A Textbook for Physicians, Veterinarians and Zoologists*, pp. 108–247. Butterworths, London.

Accepted 21 February 2019
First published online 21 March 2019